Universiteit Leiden

Psychologie
Faculteit der Sociale Wetenschappen

Differential Prognostic Value of Personality psychopathology for Treatment Outcome in Singular Anxiety disorders, Singular Mood disorders, and Comorbid Mood and Anxiety disorders

M. Shahabi

S1121480
Master Thesis Clinical Psychology
Supervisor: Prof. Dr. E. de Beurs
Institute of Psychology
Leiden University
24-08-2017

Content

Abstract

The purpose of the present study was to compare the prognostic value of personality psychopathology (PP) between patients with singular mood disorders, singular anxiety disorder, and comorbid mood and anxiety disorders. Previous studies showed that PP hampers treatment outcome more negatively in patients with mood disorders than in patients with anxiety disorders. However, differences in methodology, such as correcting for pretest severity or not, have yielded inconsistent findings between studies. A high negative impact of PP upon treatment outcome was predominantly expected for singular mood disorders and comorbid mood and anxiety disorders above and beyond the influence of pretest severity. An observational study was carried out and assessed 5755 patients with mood and/or anxiety disorders. Treatment outcome was operationalized as a continuous variable as well as a dichotomous variable. PP and symptoms of general psychopathology were assessed with the Dimensional Assessment of Personality Pathology Short Form and the Brief Symptom Inventory, respectively. Results showed a quite small negative effect of PP on outcome after controlling for pretest severity, gender, and age. The effects of PP were similar between the three diagnostic groups. Pretest severity was the best predictor of a negative outcome in all analyses. The potential confounding effects of pretest severity, and intensity and duration of treatment may explain the inconsistent findings between our study and previous studies. We advise further research to take these confounders into account when considering examination of the prognostic value of PP on treatment outcome.

Keywords: personality psychopathology, mood disorders, anxiety disorders, treatment outcome, pretest severity, intensity

1. The differential prognostic value of personality psychopathology for treatment outcome of singular disorders, singular mood disorders, and comorbid mood and anxiety disorders

1.1 Personality psychopathology

A personality disorder is described as "an enduring pattern of inner experience and behaviour that deviates markedly from expectations of the individual's culture, is pervasive and inflexible, has an onset in adolescence or early adulthood, is stable over time, and leads to distress or impairment" (DSM-5; American Psychiatric Association, 2013). The development and symptomatology of personality disorders is rather dimensional and continuous than categorical and dichotomous (Widiger, 1991), and therefore it is better to conceptualize personality disorders as personality psychopathology that represents maladaptive personality traits (Mulder, 2002). This form of psychopathology has often been dimensionally linked to axis I psychiatric disorders, such as mood and anxiety disorders (Eurelings-Bontekoe, Verheul, & Snellen, 2009; Siever, & Davis, 1991; Livesley, 1993; 1998; Joseph et al., 2001). In fact, Carlier et al. (2014) found that patients with comorbid mood and anxiety disorders displayed the most personality psychopathology (PP), followed by patients with singular mood disorders and singular anxiety disorders. They recommended to analyse the prognostic value of PP on treatment outcome for these three group of patients based on the finding that it could provide valuable information for the development of personalized treatments (Judd, Schettler, Coryell, Akiskal, & Fiedorowicz, 2013; Reich, 2007). According to clinical lore, complex clinical cases (i.e. patients with a primary anxiety or mood disorder as well as comorbid PP) have more often a disappointing treatment outcome than cases uncomplicated by comorbid PP. In view of the literature and this common opinion among therapists, the question whether the PP level has prognostic value for the treatment of mood and anxiety disorders has clinical importance.

1.2 Primary disorder

In the past, several studies have focused on analysing the prognostic value of PP for treatment outcome in different types of mood and anxiety disorders. Studies on anxiety disorders report little to no influence of PP on treatment outcome. For example, Dreessen and Arntz (1998) found little to no support for the prognostic value of PP for anxiety disorders in their review. They compared treatment response between patients with a comorbid personality disorders

and patients without a comorbid personality disorder. Treatment response was either defined as dichotomous variable that classified between responders and non-responders at end state based on whether there was a 50% reduction of symptoms or not, or as a continuous variable (i.e. posttest score corrected for pretest severity). The majority of the included studies found little to no influence of PP on treatment outcome (Chambless et al., 1995; Dreessen, Hoekstra, & Arntz, 1997; Hoffart, & Martinsen, 1993; Noyes et al., 1990). Another study by Kampman, Keijsers, Hoogduin, and Hendriks (2008) analysed treatment outcome for panic disorder, and did not find any influence of PP. On the contrary, these findings are rejected by other studies that did find a negative (and sometimes a positive) effect of PP on treatment outcome of anxiety disorders (Goddard, Wingrove, & Moran, 2015; Schat et al., 2015; Telch, Kamphuis, and Schmidt 2011; Sanatinia et al., 2016). The effect of PP was, however, small in some studies (Goddard et al., 2015; Telch et al., 2011). Treatment outcome was again either defined as a continuous variable, such as the change from pretest to posttest, or as a dichotomous variable, such as responders and non-responders. Goddard et al. (2015) examined both kind of outcomes, and found similar effects PP. Furthermore, the majority of the studies only analysed the influence of PP for one specific anxiety disorder, such as panic disorder (Noyes et al., 1990; Telch et al., 2011). Therefore, the generalizability of the prognostic value of PP to other anxiety disorders seems limited.

Additionally, studies provide more convincing evidence for a greater negative impact of PP on treatment outcome for mood disorders than for anxiety disorders. For instance, van den Hout, Brouwers, and Oomen (2016) compared the prognostic value of PP for treatment outcome between three groups of patients with anxiety disorders and one group of patients with major depression. The depressed group showed less improvement in outcome measures, while the anxiety group was not affected by the presence PP (van den Hout et al., 2016). Their findings are in line with findings from other studies showing a negative impact of PP on treatment outcome of mood disorders (Shea et al., 1990; Goddard et al., 2015). Other studies, however, do not support this stance and report that PP has little to no impact on treatment outcome (Harte, & Hawkins, 2016; Mulder, 2002). Harte et al. (2016) even found more prognostic value of PP for generalized anxiety disorder and obsessive-compulsive disorder than for depression. Nonetheless, lack of peer-reviewed instruments, small data samples, and confounding variables (i.e. pretest severity) were addressed as some of the methodological

problems that account for the inconsistency between the studies (Mulder, 2002). A systematic meta-analysis conducted by Newton-Howes et al. (2014) implemented the pooling-technique and included studies with peer-reviewed instruments, as recommended by Mulder (2002). They subsequently found a negative effect of PP on treatment outcome for mood disorders. Noteworthy, results from this meta-analysis were limited in its power due to lack of control for pretest severity (Newton-Howes et al., 2014). Furthermore, treatment outcome was either defined as a continuous outcome variable, such as posttest score corrected for pretest severity or change scores (van den Hout et al., 2016; Goddard et al., 2015), or as a dichotomous variable, such as recovered versus unrecovered at end state functioning. A patient was considered recovered when it sustained treatment response until the end of treatment (Newton-Howes et al., 2014).

Furthermore, in past research PP has also been associated with comorbid mood and anxiety disorders (Carlier et al., 2014). Studies have found lifetime comorbidity rates of 73% and 71%-85% between mood and anxiety disorders in community and clinical samples, respectively (Brown, Campbell, Lehman, Grisham, & Mancill, 2001; Lewinsohn, Zinbarg, Seeley, Lewinsohn, & Sack, 1997). Several studies showed that PP adversely effects treatment outcome when it involves patients with both mood and anxiety disorder (Tyrer, Seivewright, Ferguson, Murphy, & Johnson, 1993; Seivewright, Tyrer, & Johnson, 1998). For instance, cluster C personality disorders (i.e. avoidant, dependent, and paranoid) were found to predict worse outcomes in patients with unipolar depression and panic disorder and/or agoraphobia (Hoffart, & Martinsen, 1993). In view of the literature, it seems of clinical interest to investigate the differential prognostic value of PP in a group of patients with singular mood and anxiety disorders as well as comorbid mood and anxiety disorders.

1.3 PP, outcome measurements, and data-analysis

Whether a prognostic value is found may depend on how outcome is operationalized and measured, and what kind of data-analysis is conducted. These choices influence the findings of studies into the prognostic value of PP. In addition, findings may vary by the kind of PP and the kind of primary disorder. Thus, it is clinically relevant to search for differences between them. Furthermore, findings can also be sensitive to unmeasured confounders. For instance, Mulder (2002) recommended to control for the confounding effects of pretest severity when examining the impact of PP on treatment outcome of axis I disorders. Pretest

severity could have influenced findings from previous studies, as it is often demonstrated that pretest severity has a consistent and strong influence on posttest outcome of mood and anxiety disorders (Mulder, 2002; Kampman et al., 2008). For instance, the negative effects of PP were modest relative to the effects of pretest severity in the study by Hoffart and Martinsen (1993). Only Goddard et al. (2015), Kampman et al. (2008), Telch et al. (2011), and many studies reviewed in Dreessen and Arntz (1998) controlled for the potential effect of this confounder.

1.4 Research questions and hypotheses

Despite many research conducted into the prognostic value of PP for treatment outcome in mood and anxiety disorders, conclusive findings have not yet been found. In light of the mixed findings, the purpose of the present study was to further investigate the differential prognostic value of PP for treatment outcome in singular and comorbid mood and anxiety disorders above and beyond the influence of pretest severity. This study examined the prognostic value of PP in three different groups of patients: those with singular anxiety disorders, those with singular mood disorders, and those with comorbid mood and anxiety disorder. To analyse the prognostic value of PP independently, this study controlled for pretest severity by analysing two types of treatment outcomes: pretest severity posttest score corrected for pretest severity and pre-to-posttest change scores. Gender and age was also controlled for during the analyses. Based on the literature, we expected a higher impact of PP on treatment outcome for mood disorders as compared to anxiety disorders. Consequently, the first hypothesis was that PP has a greater negative effect on treatment outcome for singular mood disorders than for singular anxiety disorders. The second hypothesis was that PP has a greater negative effect on treatment outcome for comorbid mood and anxiety disorders than for singular anxiety disorders. Finally, this present study explored and compared the effect of PP on treatment outcome between comorbid mood and anxiety disorders and singular mood disorders.

2. Methods

2.1 Patients

The data used in this present study comprised of a sample of 5755 outpatients who received treatment at GGZ Rivierduinen or the psychiatric department of Leiden University Medical Centre between 2004 and 2013. Patients with other disorders, such as somatoform disorders

or personality disorders, were excluded from the study. The data for this study included patients with a singular anxiety disorder, a singular mood disorder, or comorbid mood and anxiety disorders. The age of the included patients ranged from 17 to 82 years old.

2.2 Research design

Data were derived from a prospective cohort study which was carried out to assess treatment outcome for patients with mood, anxiety, and somatoform disorders. For the present analysis, we used data from assessments that took place on two occasions; at the start of treatment and after six to eight months of treatment.

2.3 Setting and procedures

The first assessment took place as part of the intake procedure during which the patients were interviewed by a clinician. Patients were further asked to complete several self-report questionnaires (i.e. pretest). A second appointment was made after the interview to establish the treatment plan. Accordingly, patients were assigned to pharmacotherapy, psychotherapy (predominantly cognitive behavioural therapy), or combination therapy. At fixed time intervals patients were reassessed to monitor their progress (de Beurs, 2011). The second assessment took place at an evaluation session during which the patients were asked to again complete self-report questionnaires (i.e. posttest).

2.4 Measurement instruments

The present study assessed the following constructs: personality psychopathology as measured with the Dimensional Assessment of Personality Pathology Short Form at pretest, and severity of psychopathology as measured repeatedly with the Brief Symptom Inventory at pretest and posttest. Furthermore, the Dutch version of the Mini International Neuropsychiatric Interview-Plus was used during the intake procedure to establish the diagnosis of mood and anxiety disorders (van Vliet, & de Beurs, 2007). Psychometric evaluations demonstrated very good inter-rater and test-retest reliability with kappa coefficients between 0.76 and 1.00 (Lecrubier et al., 1997).

Personality psychopathology. The Dimensional Assessment of Personality Pathology Short Form (DAPP-SF) is a 136-item self-report questionnaire assessing maladaptive personality traits (de Beurs, Rinne, van Kampen, Verheul, & Andrea, 2009). The items are clustered into 18 subscales and four higher-order constructs. The subscales submissiveness, cognitive distortion, identity problems, affective lability, oppositionality, anxiousness,

suspiciousness, social avoidance, narcissism, insecure attachment, and self-harm underlie the first higher-order construct "Emotional Dysregulation" (ED). The subscales stimulus seeking, callousness, rejection, and conduct problems underlie the second higher-order construct "Dissocial Behaviour" (DB). The subscales intimacy problems and restricted expression underlie the third higher-order construct "Inhibitedness" (IH). Finally, the subscale compulsivity underlies the fourth higher-order construct "Compulsivity" (CO; de Beurs et al., 2009). The items are rated on a five-point Likert scale. DAPP-SF scores indicate the presence of personality pathology. Psychometric evaluations performed in community samples and clinical samples (i.e. patients with both axis-I and axis-II disorders) demonstrated good internal consistency with Cronbach's alphas between 0.76 and 0.91 (de Beurs et al., 2009). Another psychometric evaluation performed in a community sample demonstrated good convergent validity for the higher-order constructs and subscales with correlations between 0.42 and 0.71 (van Kampen, De Beurs, & Andrea, 2008). For the present statistical analysis, a two-staged procedure was chosen: if significant results were found for the higher-order constructs, underlying subscales were analysed as well (Carlier et al., 2014). Mean scores were calculated for each higher-order construct and each subscale, and were used as independent variables for the present statistical analysis.

Symptoms of general psychopathology. The BSI-53 is a 53-item questionnaire assessing symptoms of depression, anxiety, somatization, obsession-compulsion, interpersonal sensitivity, hostility, phobic anxiety, paranoid ideation, and psychoticism (Derogatis, & Melisaratos, 1983). The items are rated on a five-point Likert scale, ranging from "not-at-all" to "extremely". A psychometric evaluation was performed in a large population of psychiatric patients and demonstrated good test-retest reliability and good internal consistency with Cronbach's alphas between 0.71 and 0.84 (De Beurs, & Zitman, 2005). The BSI total score ranged from 0 to 4, and was used as a dependent variable for the present statistical analysis.

2.5 Operationalisation of outcome

Treatment outcome was operationalized as the posttest score corrected for pretest severity (i.e. residual change scores). In past research, treatment outcome has been operationalized as percentage improvement (i.e. 50% reduction of symptoms; Dreessen, & Arntz., 1998; Schat et al., 2015). However, several disadvantages such as diminished statistical power and loss of valuable information have been associated with this method, as this categorical method

classifies patients into "responders" and "non-responders" based on a cut-off of 50% (de Beurs et al., 2015; Schat et al., 2013). Conversely, residual change scores (continuous variable) have shown more statistical power than categorical variables (de Beurs et al., 2015). Furthermore, as residual change scores do not provide clinically meaningful information, an alternative method called Reliable Change Index proposed by Jacobson and Truax (1991) was applied as an adjunct to analysis with residual change scores. This method emphasizes the importance of interpreting the amount of change between pretest and posttest, therefore pre-to-posttest change scores were computed instead of residual change scores for the Reliable Change analysis. The Reliable Change Index transforms change scores into clinically more informative terms: deteriorated, unchanged, or improved (Jacobson, & Truax, 1991). To classify the patients, a cut-off score of 0.35 was chosen for the BSI change score. Patients were classified as reliably improved when their change score on the BSI was > 0.35 (a change in the direction of improvement). Patients were classified as reliably deteriorated if the change score was < -0.035 (a change in the direction of deterioration). Patients with change scores between -0.35 and 0.35 were classified as unchanged.

2.5 Statistical analysis

Statistical Package for the Social Science (SPSS version 23; de Vocht, 2013) was used for the statistical analyses, which included descriptive statistics and testing statistics. The hypotheses "whether PP has a greater negative effect on treatment outcome for mood disorders than for anxiety disorders", and "whether PP has a greater negative effect on treatment outcome of comorbid mood and anxiety disorders than for singular anxiety disorders" were tested by analysing the results of multiple regression analysis and hierarchical regression analysis. These results were further used to explore and compare the effect of PP on treatment outcome between comorbid mood and anxiety disorders and singular mood disorders.

The sequence of analysis of residual change scores was as follows: firstly, BSI pretest score, gender, and age were entered as independent variables together with the higher-order PP constructs scores in a multiple regression model with the BSI posttest score as the dependent variable. This analysis was performed for the group with singular mood disorders (depression group), the group with singular anxiety disorders (anxiety group), and the group with comorbid mood and anxiety disorders (comorbid group). Secondly, hierarchical regression analysis was conducted to analyse which higher-order construct was the best

predictor of BSI posttest score. This second analysis was conducted by choosing the stepwise entry method. The results of the multiple regression analysis and hierarchical analysis were further used to explore and compare the effect of higher-order PP constructs on BSI posttest score between comorbid mood and anxiety disorders and mood disorders. Third, if DAPP higher-order PP constructs had significant effects, the prognostic value of the 18 underlying subscales of the constructs were analysed as well. Only multiple regression analysis was conducted for this analysis. Construct CO has only one subscale, therefore custom tables were inspected to see whether the mean score of the subscale was the same as the mean score of the higher-order construct. We chose to only analyse the higher-order construct in the case the mean scores were the same. Furthermore, raw data of all independent and dependent variables were measured at an interval level. For every analysis, a two-tailed test with a significance level $p \leq 0.05$ was chosen.

The following inspections were done to meet the assumptions of multiple linear regression: the Variance Inflation Factor of the variables was analysed to rule out multicollinearity. A Variance Inflation Factor between one and ten indicates that the assumption of linear independence is not violated. When a Variance Inflation Factor (VIF) exceeds ten, it is very likely that two or more predictor variables are highly correlated with each other (de Vocht, 2013). Variable deletion is then considered in order to prevent distortion of results (Pituch, & Stevens, 2016). Additionally, the presence of influential cases within the independent variables was calculated with Cook's distance and leverage values. Cases with a Cook's distance larger than 1 and/or leverage value larger than 0.5 have substantial influence on regression coefficients (de Vocht, 2013). Scatterplots were visually inspected to further analyse such cases. In addition, P-P plots and scatterplots were created to control if residuals were normally distributed, and if there was homogeneity of variance (i.e. homoscedasticity). Finally, bivariate correlations were conducted to analyse the relationship between BSI posttest score and each higher-order construct.

The sequence of analysis of change scores according to the Reliable Change Index (RCI; Jacobson, & Truax, 1991) was as follows: first, Delta BSI score was computed by subtracting the pretest score from the posttest score on each measure. Second, Reliable Change variable was computed based on the Delta BSI score. Reliable Change is a categorical variable with three levels: improved (Delta BSI score > 0.35), unreliable change (Delta BSI

score ≤ 0.35 and ≥ -0.35) and deteriorated (Delta BSI score < -0.35). Reliable Change was recoded into two dummy variables: Reliable Improvement (1 = improved, 0 = unchanged), and Reliable Deterioration (1= deteriorated, 0 = unchanged). Binary logistic analysis was conducted to analyse whether the four higher-order PP constructs predict categorical outcome for all three diagnostic groups. Again, two-tailed tests with a significance level $p \leq 0.05$ were chosen.

3. Results

3.1 Sample characteristics

The data sample of *N* = 5755 included patients with singular anxiety disorders, singular mood disorders, and comorbid mood and anxiety disorders. The mean age was M = 3893, SD = 12.5, and most of the patients were female (64.2%). With regard to clinical data, most patients were diagnosed with comorbid mood and anxiety disorders (n = 2061; 26.4%) patients). A further *N* = 2014 (25.8%) were diagnosed with a singular anxiety disorder. Finally, *N* = 1680 (21.5%) were diagnosed with a singular mood disorder. Demographic data of the sample are presented in Table 1. Frequencies of specific mood and anxiety disorders are presented in Appendix 1. Major depressive disorder (recurrent, moderate), posttraumatic stress disorder, and generalized anxiety disorders were the most common disorders in the data sample: *N* = 1232 (21,4%), *N* = 898 (15,6%), and *N* = 898 (15,6%), respectively.

Table 1

Baseline characteristics in 5755 outpatients diagnosed with mood or anxiety disorders, who received treatment at GGZ Rivierduinen or the LUMC between 2004 and 2013

Demographic data	M(SD)	
Age	38.9(12.5)	
	N	%
Gender		
Male	2138	37,2%
Female	3617	62,8%

3.2 Preliminary analyses

The total number of missing values on the DAPP self-report questionnaire scores was only

four. The CO construct had three missing values, and the DB construct had one missing value. These four cases were excluded from the regression analysis. As the number of missing values was low, no further analysis of missing values was conducted. Predictor variables had a Variance Inflation Factor between one and ten. This indicates that the variables were sufficiently uncorrelated. The P-P plot indicated that errors were normally distributed between independent and dependent variables. Some points were not completely on the line, but close. The scatterplot of standardized residuals indicated that the assumption of homogeneity of variance was met. Cook's distance and leverage values demonstrated that there were no individual cases with large influence on the regression coefficients. The highest Cook's distance values were 0.016, 0.018, and 0.14 for the depression group, anxiety group, and comorbid group, respectively. The highest leverage values were 0.017, 0.014, 0.022 for the three groups, respectively. With regard to bivariate correlations, the relationship between BSI pretest score and BSI posttest score was significant and of medium size, $r = 0.540$, $p < 0.05$. The relationship between gender and BSI posttest score was not significant, $r = 0.003$, $p = 0.829$. Finally, the relationship between age and BSI posttest score was significant and weak, $r = -0.030$, $p < 0.05$. After controlling for BSI pretest score, gender, and age, only BSI pretest score remained significant in multiple regression models. BSI pretest score was also the best predictor of BSI posttest score in hierarchical regression models for all three groups. It had B-values of 0.491, 0.453 and 0.351 for comorbid group, depression group, and anxiety group, respectively (Table 2). Within the depression group, 25% of the variance was explained by BSI pretest score. Finally, 26% and 20% of the variance was explained by BSI pretest score within the anxiety group and comorbid group, respectively. Gender and age were not significant predictors in the hierarchical models for anxiety and comorbid group. Only gender was significantly associated with outcome in the depression group: woman had better outcomes than men.

3.3 Does PP have a greater negative effect on treatment outcome for singular mood disorders than for singular anxiety disorders?

Next, the four higher-order PP constructs (ED, DB, IH, and CO) were analysed in addition to BSI pretest score, gender, and age. Table 2 displays the effect of the constructs on the BSI posttest score for the depression group, the anxiety group and the comorbid group.

The findings for the singular depression group were as follows: higher ED and IH scores, and lower CO scores independently predicted higher BSI posttest scores after controlling for BSI pretest score, gender, and age. DB was not a significant predictor of BSI posttest score. The model was significant and had a medium R-square value of $R^2 = 0{,}257$, $p < 0.05$. After applying the hierarchical method with stepwise entry, higher ED appeared to predict BSI posttest score, followed by higher IH and lower CO. DB remained a nonsignificant predictor of BSI posttest score. The addition of the significant higher-order PP constructs was associated with small, significant improvements in $R^2 = 0{,}253$, $p < 0.05$; ED = 0.007; IH = 0.002, CO = 0.002.

Only Higher ED scores independently predicted higher BSI posttest scores within the anxiety group after controlling for BSI pretest score, gender, and age. Higher DB and IH scores were significant in the multiple regression analysis, but did not remain significant in the hierarchical analysis. CO was not a significant predictor of BSI posttest score in both analyses. The model of the multiple regression analysis was significant and had a medium R-square value of $R^2 = 0{,}287$, $p < 0.05$. Similar to the depression group, the significant higher-order ED construct only had small significant changes in the hierarchical model, $R^2 = 0{,}284$, $p < 0.05$; ED = 0,027.

In summary, the higher-order PP constructs did not have a greater negative effect on treatment outcome for mood disorders than for anxiety disorders. The constructs were significant in the multiple regression model, but explained very little additional variance in the hierarchical model for both groups. ED was the best predictor of treatment outcome for both groups, and explained more variance within the anxiety group than within the depression group. The negative effects of the remaining higher-order PP constructs on treatment outcome for the depression group were similar to the effects on treatment outcome for the anxiety group. CO had a significant positive effect on treatment outcome for depression group.

3.4 Does PP have a greater negative effect on treatment outcome for comorbid mood and anxiety disorders than for singular anxiety disorders?

Similar to the results for the depression group, higher ED and higher IH scores, and lower CO scores independently predicted higher BSI posttest scores after controlling for BSI pretest score, gender, and age. DB was not a significant predictor of BSI posttest score. The model

was significant and had a medium R-square value of R^2 = 0,230, $p < 0.05$. After applying the hierarchical method with stepwise entry, higher IH appeared to predict BSI posttest score the most, followed by higher ED and lower CO. Similar to the results of the singular groups, the significant predictors only had small, significant changes in R^2 = 0,230, $p < 0.05$; IH = 0.006; ED = 0.001; CO = 0.002. Again, DB was not a significant predictor in the hierarchical model.

In summary, the higher-order PP constructs did not have a greater negative effect on treatment outcome for the comorbid group than for the anxiety group. ED also explained more variance within the anxiety group than within the comorbid group. The negative effects of the remaining higher-order PP constructs on treatment outcome for the comorbid group were similar to the effects on treatment outcome for the anxiety group. CO had a significant positive effect on treatment outcome for comorbid group.

3.5 Explorative analysis of the difference in prognostic value of PP between comorbid mood and anxiety disorders and singular mood disorders

Table 2 shows that the effect of the higher-order PP constructs on BSI posttest score was relatively similar between the comorbid group and the depression group. The biggest difference in the effect of PP was a difference of 0.042 for IH; B-values of IH were 0.058 and 0.100 for the depression and the comorbid group, respectively. The amount of explained variance was similar and small: R^2 = 0.257 and R^2 = 0.230 for the depression group and comorbid group, respectively. ED had a slightly greater negative effect on BSI posttest score for the depression group than for the comorbid group. IH and CO had a slightly greater negative effect on BSI posttest score for the comorbid group than for the depression group.

Table 2

Multiple regression analysis examining the effect of higher-order PP constructs on BSI posttest score after a maximum of six months of treatment

	B (SE)	Beta	*T*	*P*	95% CI
Selection variable: depression group					
F(7, 1671), 82,48, $p < 0.05$; Adj. R^2 = 0.254					
ED	0,091(0,031)	0,087	2,904	0,004**	0,030-0,153
IH	0,058(0,023)	0,056	2,477	0,013**	0,012-0,104

CO	-0,033(0,016)	-0,046	-2,092	0,037**	-0,064 to -0,002
DB	0,049(0,030)	0,042	1,622	0,105	-0,010 to 0,108
BSI pretest score	0,453(0,028)	0,431	16,064	0,000**	0,397-0,508
Gender	-0,035(0,031)	-0,026	-1,125	0,261	-0,096 to 0,026
Age	0,001(0,001)	0,024	1,070	0,285	-0,001 to 0,003
Selection variable: anxiety group					
$F(7, 2004) = 115{,}49$, $p < 0.001$; Adj. $R^2 = 0.287$					
ED	0,178(0,025)	0,200	7,018	0,000**	0,128-0,228
DB	0,051(0,025)	0,045	2,042	0,041**	0,002-0,101
IH	0,038(0,018)	0,042	2,082	0,037**	0,002-0,075
CO	-0,019(0,012)	-0,031	-1,541	0,124	-0,042-0,005
BSI pretest score	0,351(0,028)	0,363	14,380	0,000**	0,303-0,398
Gender	0,010(0,024)	0,009	0,436	0,663	-0,036 to 0,057
Age	0,002(0,001)	0,038	1,941	0,052	0,000-0,003
Selection variable: comorbid anxiety and depression group					
$F(7, 2052) = 87{,}61$, $p < 0.05$; Adj. $R^2 = 0.230$					
ED	0,071(0,035)	0,055	2,028	0,043**	0,002-0,140
IH	0,100(0,027)	0,079	3,756	0,000**	0,048-0,152
CO	-0,038(0,017)	-0,045	-2,210	0,027**	-0,072 to -0,004
DB	0,024(0,034)	0,016	0,695	0,487	-0,043 to 0.090
BSI pretest score	0,491(0,029)	0,421	16,801	0,000**	0,433-0,548
Gender	0,000(0,036)	0,001	0,003	0,998	-0,071 to

Age	0,000(0,001)	0,003	0,171	0,864	0,071 -0,003 to 0,003

Note: ED denotes Emotional Dysregulation; IH denotes Inhibitedness; CO denotes Compulsivity; DB denotes Dissocial behaviour; CI denotes confidence interval

* adjusted for gender and age

** $p < 0.05$

3.6 Unique subscales of the higher-order constructs

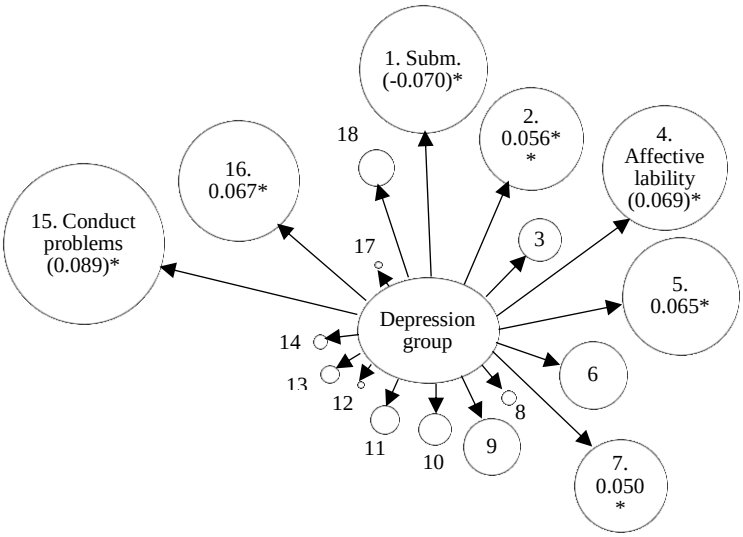
Next, the unique subscales of the significant higher-order PP constructs were analysed in addition to pretest BSI score, gender, and age for all three groups. The subscales of ED were analysed within all three groups. The subscales of IH were only analysed within the depression and comorbid group. This higher-order construct was not a significant predictor of treatment outcome within the anxiety group. The same holds for DB whose subscales were only analysed within the anxiety group. Subscale CO was not analysed, as it had the same mean value as its higher-order construct CO; M = 2.89, M = 2.87, and M = 3.01 for the anxiety group, depression group, and comorbid group, respectively. Predictor variables had a Variance Inflation Factor between one and ten. This indicates that the variables were sufficiently uncorrelated. Cook's distance and leverage values demonstrated that there are no individual cases with high influence on the regression coefficients. The highest Cook's distances were 0.028, 0.019, and 0.009 for the depression group, anxiety group, and comorbid group, respectively. The highest leverage values were 0.040, 0.44, and 0.44 for the three groups, respectively.

The findings for the depression group were as follows: higher cognitive distortion, affective lability, oppositionality, suspiciousness, and lower submissiveness scores (all subscales of ED), and higher intimacy problems (a subscale of IH) scores independently predicted a higher BSI posttest score after controlling for BSI pretest score, gender, and age. These analyses resulted in significant and medium R-square values of $R^2 = 0{,}271$ and the $R^2 = 0.248$, $p < 0.05$ for the ED model and IH model, respectively.

Similarly, higher Cognitive Distortion, Oppositionality, Suspiciousness, Self-harm, and Lower submissiveness scores (all subscales of ED) independently predicted higher BSI posttest score for the anxiety group after controlling for BSI pretest score, gender, and age. The ED model had a significant and medium R-square value of $R^2 = 0.299$, $p < 0.05$.

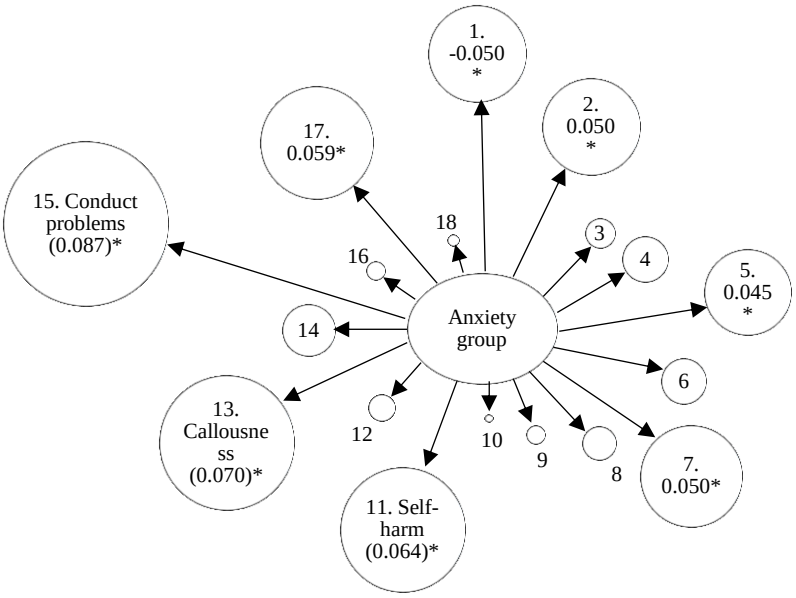
The significant subscales of ED and IH for the depression group with the addition of higher Self-harm and lower Narcissism also independently predicted independently predicted higher BSI posttest score for the comorbid group after controlling for BSI pretest score, gender, and age (except for oppositionality and cognitive distortion). Likewise, significant and medium R-square values of $R^2 = 0.238$ and $R^2 = 0.231$, $p < 0.05$ were found for the ED model and the IH model, respectively.

Based on the results regarding the subscales three figures were created, as shown below, where the figures represent the three groups. The size of the circle corresponds to the amount of influence that the respective variable exerts. All 18 subscales instead of only the significant subscales were included to demonstrate the proportionality between significant subscales and nonsignificant subscales. As such, the smallest circles were not significant and had very small B-values ranging from 0.001 to 0.036. The t values, *p* values and confidence intervals in addition to the prognostic values of the subscales for all three groups are presented in Appendix 2.



Legenda

1. Submissiveness
2. Cognitive distortion
3. Identity problems
4. Affective lability
5. Oppositionality
6. Anxiousness
7. Suspiciousness
8. Social avoidance
9. Narcissism
10. Insecure attachment
11. Self-harm
12. Stimulus seeking
13. Callousness
14. Rejection
15. Conduct problems
16. Intimacy problems
17. Restricted expression
18. Compulsivity



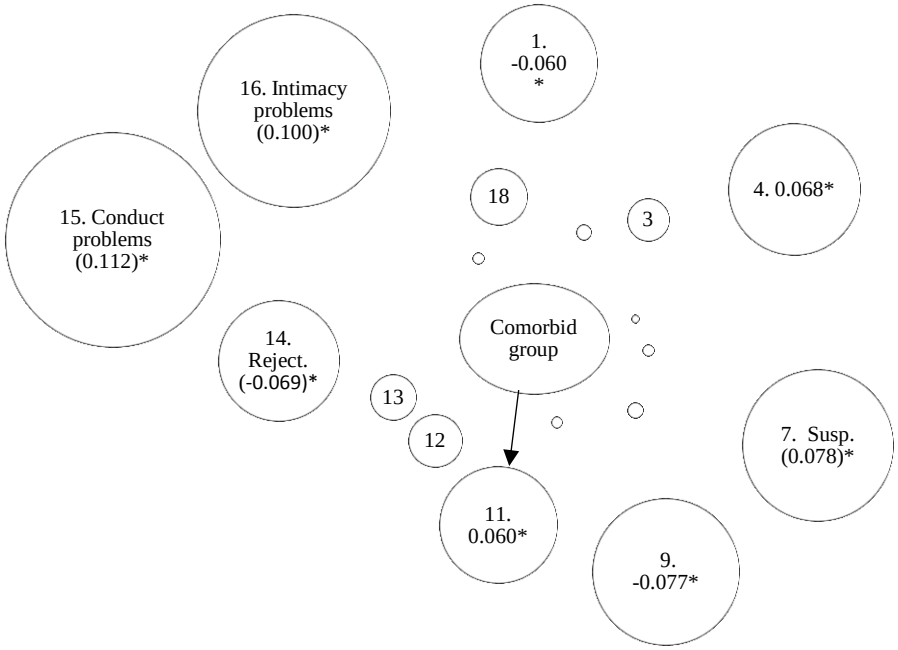


Figure 1.

Prognostic profile based on the B-value of the 18 unique subscales of the four higher-order PP constructs for depression, anxiety, and comorbid group. * $p < 0.5$

3.7 Reliable Change Index according to Jacobson and Truax (1991)

The Reliable Change Index showed that 3205(55.7%) patients were reliably improved with a Delta BSI score greater than 0.35. A further N = 2011(34.9%) patients had no reliable change. Finally, 539(9.4%) patients were reliably deteriorated with a Delta BSI score less than -0.35.

Next, the association between higher-order PP constructs and Reliable Improvement was compared between the three groups of patients. The findings for the depression group were as follows: a test of the full model for Reliable Improvement was statistically significant, $\chi^2(6, N = 1680) = 43.34$, $p < 0.001$. The model had an overall success rate of 64.9%. Only ED and CO had significant partial effects. Moreover, the model fitted the data as Hosmer and Lemeshow was not significant, $p = 0.089$. The positive value for both higher-order PP constructs indicated that patients who had high ED and CO scores were more likely to be in the improved group than in the unchanged group.

The findings for the anxiety group were as follows: a test of the full model for Reliable Improvement was statistically significant, $\chi^2(6, N = 2014) = 119.33$, $p < 0.001$. The model had an overall success rate of 61.5%. Only ED and CO had significant partial effects. Hosmer and Lemeshow was not significant, $p = 0.617$. Similar to depressed patients, patients with anxiety disorders and high ED and CO scores were more likely to be in the improved group than in the unchanged group.

The findings for the comorbid anxiety and depression group were as follows: a test of the full model for Reliable Improvement was statistically significant, $\chi^2(6, N = 2061) = 23{,}71$, $p < 0.001$. The model had an overall success rate of 66.9%. Only ED had significant partial effects. Hosmer and Lemeshow was not significant, $p = 0.23$. The positive value for ED indicates that patients who have a comorbid mood and anxiety disorder as well as high ED scores are more likely to be in the improved group than in the unchanged group. Table 3 on the next page displays the logistic coefficients, Wald test, and odds ratio of each of the predictors for the three groups.

Table 3

Binary logistic regression predicting Reliable Improvement from higher-order PP constructs within depression group, anxiety group, and comorbid group

Predictor	B(SE)	Wald (χ^2)	P	Odds ratio	95% CI for odds ratio
Selection variable: depression group					
ED	0.439(0.106)	17.245	0.000**	1.551	1.26-1.90
DB	-0.165(0.119)	1.930	0.165	0.848	0.67-1.07
IH	-0.005(0.091)	0.003	0.955	0.995	0.83-1.19
CO	0.136(0.062)	4.891	0.027**	1.146	1.02-1.29
Gender	0.211(0.120)	3.073	0.080	1.234	0.98-1.56
Age	-0.009(0.004)	3.881	0.049	0.991	0.98-1.00
Selection variable: anxiety group					
ED	0.784(0.095)	61.439	0.000**	2.114	1.75-2.55
DB	-0.038(0.113)	0.115	0.734	0.962	0.77-1.20
IH	0.003(0.083)	0.001	0.976	1.003	0.85-1.18
CO	0.126(0.055)	5.339	0.021**	1.135	1.02-0.26
Gender	0.064(0.107)	0.357	0.550	1.066	0.86-1.32
Age	-0.005(0.004)	1.522	0.217	0.995	0.99-1.00
Selection variable: comorbid anxiety and depression group					
ED	0.310(0.095)	10.59	0.001**	0.73	0.61-0.88
DB	0.022(0.107)	0.04	0.838	1.02	0.83-1.26
IH	0.088(0.083)	1.12	0.289	1.09	0.93-1.29
CO	-0.103(0.055)	3.49	0.062	0.90	0.81-1.01
Gender	-0.013(0.114)	0.01	0.907	0.99	0.79-1.23
Age	0.005(0.004)	1.29	0.256	1.01	1-1.01

Note: ED denotes Emotional Dysregulation; IH denotes Inhibitedness; CO denotes Compulsivity; DB denotes Dissocial behaviour; CI denotes confidence interval

* adjusted for gender and age

** $p < 0.05$

Next, the association between higher-order PP constructs and Reliable Deterioration was compared between the three groups of patients. The full model for Reliably Deterioration was statistically not significant, $\chi^2(6, N = 2014) = 8.642$, $p = 0.195$ and $\chi^2(6, N = 2061) = 5.93$, $p = 0.431$ for the depression group and comorbid group, respectively. The four higher-order constructs were not further examined within these two groups. The full model for Reliable Deterioration was statistically significant for the anxiety group, $\chi^2(6, N = 2014) = 29.75$, $p < 0.001$. The model had an overall success rate of 83.8%. Only ED had significant partial effects. Moreover, the model fits the data as Hosmer and Lemeshow was not significant, $p = 0.74$. The positive value for ED indicated that patients who had high ED scores were more likely to be in the deteriorated group than in the unchanged group. The logistic regression coefficients, Wald test, and odds ratio of each of the predictors for the anxiety group are presented in table 4.

Table 4

Binary logistic regression predicting Reliable Deterioration from higher-order PP constructs within the anxiety group

Predictor	B(SE)	Wald (χ^2)	P	Odds ratio	95% CI for odds ratio
Emotional dysregulation	0.616(0.165)	13.996	0.000**	1.851	1.34-2.56
Dissocial behaviour	0.303(0.198)	2.330	0.127	1.354	0.92-2.00
Inhibitedness	-0.008(0.149)	0.003	0.959	0.992	0.74-1.33
Compulsivity	-0.027(0.100)	0.074	0.786	0.973	0.80-1.18
Gender	0.091(0.190)	0.230	0.631	1.095	0.75-1.59
Age	0.004(0.007)	0.282	0.595	1.004	0.99-1.02

Note: DAPP-SF denotes Dimensional Assessment of Personality Psychopathology-Short Form; CI denotes confidence interval

* adjusted for gender and age

** $p < 0.05$

3.8 Summary of results

BSI pretest score was a significant predictor of BSI posttest score within all three diagnostic

groups. The prognostic value of BSI pretest score was substantially higher than the prognostic values of the four DAPP higher-order constructs and the 18 unique subscales. Gender was also a significant predictor of BSI posttest score within the depression group, but showed little prognostic value. The amount of explained variance by these four constructs was uniform across the groups. The four higher-order constructs ED, DB, IH, and CO did not have a greater negative effect on BSI posttest score for mood disorders and comorbid mood and anxiety disorders than for singular anxiety disorders. Emotional dysregulation had a greater negative effect on BSI posttest score for the anxiety group than for the depression and comorbid group. The remainder of the constructs had relatively similar B-values between the three groups. For example, Dissocial Behaviour was not significant in all three groups. Explorative analysis of the effect on BSI posttest score between the comorbid group and the depression group showed that the negative effect of higher-order constructs on BSI posttest score were relatively similar and did not have much comparative value.

The results with regard to the 18 unique subscales of the DAPP higher-order constructs yielded a personality profile for each subgroup. Lower submissiveness scores significantly predicted higher BSI posttest scores within all three groups. Lower narcissism and rejection scores also significantly predicted higher BSI posttest score within the comorbid group. The remainder of significant subscales had negative B-values.

The results of the binary analysis according to the RCI method showed that higher ED scores predict improvement within all three groups. Higher ED scores also predicted deterioration within the anxiety group. Finally, higher CO scores predicted improvement within the depression group. IH and DB were not significant predictors of improvement and deterioration.

4. Discussion

This present study aimed to investigate whether the impact of PP on treatment outcome differs according to type of primary disorder. We compared the effects of PP on treatment outcome between three diagnostic groups: patients with mood disorders, patients with anxiety disorders, and patients with both mood and anxiety disorders. It was expected that PP would have a greater negative impact on outcome in patients with singular mood disorders or comorbid mood and anxiety disorders than for singular anxiety disorders. We also compared

the effect of PP on treatment outcome between singular mood disorders and comorbid mood and anxiety disorders. As expected, PP predicted treatment outcome above and beyond the influence of pretest severity, gender, and age. It was justified to control for pretest severity, as this demonstrated to be a consistent and important predictor of treatment outcome for all three groups. The prognostic value of pretest severity was even greater than the prognostic value of PP (i.e. the four higher-order constructs and the 18 subscales). Our finding that 20% to 26% of the variance in outcome was explained by pretest severity replicates findings of a study by Kampman et al. (2008) who found explained variance rates between 20% to 51%. It is, however, a common finding that pretest severity affects posttest treatment outcome and several studies emphasize the importance of taking this confounder into account (Mulder, 2002; Ramnerö, & Ost, 2004). Gender was also a predictor of treatment outcome, but solely for the depression group and showed negligible prognostic value.

The findings of this present study do not support the hypotheses that PP has a stronger negative impact on outcome in patients with singular mood disorders and comorbid mood and anxiety than in patients with singular anxiety disorders. Although personality psychopathology was associated with worse outcomes in all three diagnostic groups, the effect of PP on outcome was quite small. It is surprising that the negative effects of the four higher-order PP constructs and 18 subscales on treatment outcome were largely similar between the three groups. Another unexpected finding was that Emotional Dysregulation predicted worse outcome and more so for patients with anxiety disorders than for patients from the other two groups. Furthermore, we found that the higher-order PP constructs had similar negative effects on treatment outcome in patients with singular mood disorders and comorbid mood and anxiety disorders. In contrast to our hypotheses, we found that one higher-order construct (Compulsivity) was associated with better outcomes in patients with a mood disorder or a comorbid mood and anxiety disorder. Compulsivity had no prognostic value for singular anxiety disorders. At subscale level, Submissiveness was found to predict better outcomes in all three groups. Narcissism was found to predict better outcomes in patients with comorbid mood and anxiety disorders. The positive findings regarding anxiety disorders confirm an earlier finding that patients with mild to moderate personality problems and hypochondria show improvement in CBT (Sanatinia et al., 2016).

Personality psychopathology was further examined in relation to the clinical end state of a patients (i.e. improved, unchanged or deteriorated). The effect of PP on the direction of change within the patients was very similar between the diagnostic groups. Emotional Dysregulation predicted improvement on the binary outcome variable in all three groups, which is surprising given the previous finding that Emotional Dysregulation predicts a negative outcome on the continuous variable. Several reasons may explain this inconsistency. Firstly, we transformed our initial continuous outcome (i.e. residual change scores) into a dichotomous variable (end state based on change scores) to determine the direction of change, hence used two different treatment outcomes. It is possible that the effect changed according to the kind of treatment outcome. Secondly, the negative effects of the higher-order PP constructs on the continuous outcome were quite small. Due to its small magnitude, this effect may have easily changed from a negative effect into a positive effect on the binary outcome variable. Furthermore, Compulsivity predicted improvement, but solely for the depression group. This finding is in line with the results of the analyses on the continuous outcome, as this higher-order construct was found to predict a better outcome within the depression group. Apparently, Compulsivity has a positive, albeit small, effect on the outcome of treatment for depression, but not when (comorbid) anxiety is present. Perhaps that this type of PP plays a role in facilitating treatment for patients with a mood disorder. Depressed patients typically feel worthless and don't have much confidence, and some compulsivity may help them abide more closely to therapeutic instructions and assignments (Faber, & O'Guinn, 1989).

4.1 Strengths and weaknesses

This present study has high external validity due to the broad inclusion criteria, dimensional measuring of PP, and large sample size. The observational nature of the study, the fact that we controlled for pretest severity, and the fact that we examined two different treatment outcomes also strengthen our findings. By collecting data in everyday clinical practice, a large representative sample from clinical institutions was created. Hereby, the generalizability of the findings to day-to-day patients is high (Mulder, 2002). In addition, the choice to measure PP dimensionally further increased the strength of our study (Newton-Howes et al., 2014). Previous studies consistently criticized categorical definitions of PP (i.e. personality disorder), because there is still no consensus on how to classify patients with personality problems the best (Newton-Howes et al., 2014; Bernstein, Iscan, & Maser, 2007). The DAPP-SF used in

this present study measures "personality traits over a continuum of adaptive to maladaptive, whereas personality disorders are maladaptive by definition" (Carlier et al., 2014). Nevertheless, one must be aware that simply having personality traits does not equal having a personality disorder (Widiger, & Costa, 2012). The new DSM-5 takes this notion into account and developed an alternative model that proposes that the combination of diagnosed characteristics of a personality disorder and maladaptive personality traits lead to a diagnosis of personality disorder (Berghuis Kamphuis, & Verheul, 2014).

The findings should also be considered in light of several of limitations. Firstly, systematic information about the intensity and duration of the treatment was lacking. The fact that we only found little prognostic value of PP can be due to differences in treatment provided to patients with and without PP. Possibly, patients with (severe) maladaptive personality traits were treated more intensively and longer than patients without maladaptive personality traits. Ultimately, well adapted treatment may have mitigated the effect of PP on treatment outcome. A study that supports this notion is a study by Dreessen, Hoekstra, and Arntz (1997) who found no differences between PD and non-PD patients after additional treatment for the PD patients. Secondly, we do not have information about type of treatment, therefore it was not possible to control for the potential confounding effects of this factor. Shea et al. (1998) demonstrated that type of treatment has differential effects on outcome between depressed patients with PD and without PD. For example, patients with PD benefited more from cognitive behavioural therapy, whereas patients without PD benefited more from interpersonal therapy or medical therapy (Shea et al., 1998). Thirdly, we used self-report questionnaires to measure PP and the level of depression and anxiety. This type of instrument is sensitive to social desirability and requires a certain amount of self-awareness in patients for them to report accurately about their symptoms and behaviours (Edwards, 1957). However, both BSI-53 and DAPP-SF demonstrated good psychometric properties (de Beurs, & Zitman, 2005; de Beurs et al., 2009). Moreover, the DAPP-SF has proven itself to be a good measure of maladaptive personality profiles, and adequately distinguishes between samples with and without personality disorders (de Beurs, Rinne, van Kampen, Verheul, & Andrea, 2010). Fourth, the extent to which this present study can be compared to previous studies is limited, as there is no consistent methodology between the studies. This is not necessarily a limitation of the present study, but rather a general weakness of research

conducted into the prognostic value of PP. As noted in the introduction, whether a prognostic value of PP is found may depend on the operationalization of PP, the operationalization and measurement of outcome, type of primary disorder, and kind of data-analysis. Previous studies used various operationalisations of treatment outcomes and various approaches to measure PP. In addition, some studies controlled for pretest severity, whereas other studies did not.

4.2 Depression and anxiety

To an extent our findings replicate findings of previous studies that PP has a negative impact upon treatment outcome in patients with mood disorders (Shea et al., 1998; Goddard et al., 2015; van den Hout et al., 2016). The effect of PP on treatment outcome was, however, quite small as compared to what a thoroughly conducted meta-analysis by Newton-Howes et al. (2014) had found. They demonstrated convincing evidence that depressed patients with a comorbid personality disorder were twice as likely to be unresponsive to treatment as depressed patients without a comorbid personality disorder (Newton-Howes et al., 2014). This inconsistency is also found when comparing the influence of PP on treatment outcome for anxiety disorders between our study and previous studies. Remarkably, our finding that PP also had a negative prognostic value upon treatment outcome of anxiety disorders contradicts earlier conclusions that PP has no impact for anxiety disorders (Dreessen, & Arntz, 1998; Kampman et al., 2008), and replicates findings of other studies with patients with anxiety disorders (Telch et al., 2011; Schat et al., 2014; Goddard et al., 2015).

Based on the earlier notion that the prognostic value of PP differs according to type of methodology, we may assume that two factors have caused this inconsistency between our study and previous studies. Firstly, many studies did not take the importance of pretest severity into account, and found high effect of PP on treatment outcome for mood disorders (Newton-Howes et al., 2014; Shea et al., 1998), and little to no effect of PP on treatment outcome for anxiety disorders (Black, Wesner, Gabel, Bowers, & Monahan, 1994; Dreessen, Hoekstra, & Arntz, 1997). Since pretest severity proved to be a strong and consistent influence on treatment outcome in the present study and previous studies, it is plausible that PP has lesser prognostic value for mood disorders when pretest severity is controlled for. This may explain why the effect PP was small in our study compared to the high effects in studies that did not control for pretest severity (Newton-Howes et al., 2014). Likewise, previous

studies that controlled for pretest severity also found a small effect of PP on treatment outcome for both singular mood disorders. For instance, Goddard et al. (2015) found very small changes in explained variance within both the depression group and anxiety group (R-squared change of 0.005) after adding PP and pretest severity as independent variables. Similarly, van den Hout et al. (2016) found a relatively weak influence of PP (Eta square of 0.12 for group x time interaction on depression score) on treatment outcome for mood disorders.

At the same time, the influence of pretest severity could have obscured possible negative effects of PP on treatment outcome for anxiety disorders in previous studies (Black et al., 1994). For instance, Black et al. (1994) that did not control for pretest severity and found no effect of PP on treatment outcome for panic disorders. Another study by Schat et al. (2015) examined the effect of PP on treatment outcome with the DAPP-SF without controlling for pretest severity, and found that only two subscales (i.e. affective lability and behaviour problems) had a small effect on treatment outcome. The present study, on the other hand, found prognostic value for several higher-order constructs and subscales of the DAPP-SF. Nonetheless, controlling for pretest severity still did not lead to higher effects of PP on treatment outcome. As mentioned earlier, pretest severity was the best predictor of outcome in our study. Similarly, Hoffart and Martinsen found that pretest severity had a greater negative effect on treatment outcome (explained variance rates between 12.4% and 24%) than the cluster C personality disorders (rates between 4.8% and 14.9%) in patients with unipolar depression and panic disorder and/or agoraphobia. Other studies that controlled for pretest severity also found limited of PP on treatment outcome for anxiety disorders (Dreessen et al., 1997; Chambless et al., 1995; Goddard et al., 2015; Telch et al., 2011). A study that examined panic disorders even found no effect of PP on treatment outcome after controlling for pretest severity (Kampman et al., 2008).

Secondly, we highlight the intensity and duration of treatment as an important factor that may have altered the outcome in our study and previous studies. As mentioned earlier, we did not consider the potential confounding effects of this factor, and therefore we did not have an entirely pure measure of the effect of PP on treatment outcome. Some studies that also did not consider the intensity of treatment found great negative effects of PP on treatment outcome for mood disorders (Newton-Howes et al., 2014). As such, it remains unclear

whether treatment was not intensive enough for these group of patients. These findings are in contrast with many previous studies that did consider this factor by choosing a standard length of treatment or controlling for the number of sessions (Dreessen, & Arntz, 1998; Goddard et al., 2015; Kampman et al., 2008; Shea et al., 1998). For example, Goddard et al. (2015) still found a negative effect of PP after controlling for the number of sessions. Nevertheless, the intensity and duration of treatment might be an important confounder to consider in future research. A study by Steketee (1990), who treated patients with obsessive-compulsive disorder and comorbid personality disorders during an intensive behavioural treatment, acknowledges this notion and showed that the effort of the therapist can determine the treatment outcome. Steketee (1990) emphasized that the powerful effects of intensive treatment may have resulted into no differences between PD and non-PD patients regarding treatment outcome.

4.3 Implications

This present study has broadened the way on how to examine the prognostic value of PP for treatment outcome in mood and anxiety disorders by viewing PP in a dimensional manner. Only a few studies have done this before (Schat et al., 2015). The majority of studies merely focused on the presence of a comorbid personality disorder. Our study provides evidence that there is a group of patients with mood or anxiety disorders who display maladaptive personality traits without necessarily meeting diagnostic criteria for a full-blown personality disorder (Carlier et al., 2013). The possible influence of dimensional PP rather than personality disorders on treatment outcome has increasingly been acknowledged in recent years (Reich, Russel, & Vasile, 1993; Reich, 2003). Furthermore, the substantial prognostic value of pretest severity for treatment outcome in our study and previous studies demonstrates that this an important factor to consider when analysing the independent influence of PP.

The discussed findings and weaknesses also highlight one clinical implication: treatment effectiveness may not be adversely affected by the presence of maladaptive personality traits, as we only found small negative effects. Similar to our other findings, some studies even found a positive influence of some adverse personality traits, such as cluster A traits (i.e. odd, eccentric; Illardi, Craighead, & Evans, 1997). However, it is not yet certain what the pure role of PP on treatment outcome is in patients with mood and/or anxiety disorders, as pretest severity and variations in the intensity of treatment have been found to

greatly influence treatment outcome as well (Steketee, 1990; Kampman et al., 2008).

4.4 Recommendations

The discussed limitations and implications lead to two recommendations. We recommend taking the potential confounding effects of pretest severity into account when conducting further research into the prognostic value of PP for treatment outcome in mood and anxiety disorders. The fact that our study and previous studies found much more prognostic value for pretest severity than for PP supports this recommendation. Finally, the intensity, duration, and type of treatment must also be controlled for in order to purely detect the independent influence of PP on treatment outcome, as more intensive or lengthy treatments may obscure negative effects of PP on the progress of treatment and its outcome.

4.5 Conclusion

This prospective cohort study examined the differential effects of PP on treatment outcome in singular and comorbid mood and anxiety disorders among adult outpatients treated at GGZ Rivierduinen or Leiden University Medical Centre. The present findings do not confirm results from previous studies. A small effect of PP on treatment outcome was found, and demonstrated to be similar between the three diagnostic groups. Several factors such as treatment characteristics (i.e. intensity, duration, modality) and pretest severity may account for the inconsistent findings between studies regarding the prognostic value of PP. Future research is required to purely detect the independent influence of PP on treatment outcome for mood and anxiety disorders.

References

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (5th ed.)*. Washington, DC: Author.

Berghuis, H., Kamphuis, J. H., Verheul, R. (2014). Specific personality traits and general personality dysfunction as predictors of the presence and severity of personality disorders in a clinical sample. *Journal of Personality Assessment, 96,* 410-416.

Bernstein, D. P., Iscan, C., & Maser, J. (2007). Opinions of personality disorder experts regarding the DSM-IV personality disorders classification system. *Journal of Personality Disorders, 21*, 536-551.

Black, D. W., Wesner, R. B., Gabel, J., Bowers, W., & Monahan, P. (1994). Predictors of short-term treatment response in 66 patients with panic disorder. *Journal of Affective Disorders, 30,* 233-41.

Brown, A., Campbell, L. A., Lehman, C. L., Grisham, J. R., & Mancill, R. B. (2001). Current and Lifetime Comorbidity of the DSM-IV Anxiety and Mood Disorders in a Large Clinical Sample. *Journal of Abnormal Psychology, 110*, 585-599.

Carlier, I. V., Colijn, S., van Rood, Y. R., Streevelaar, M. F., van Vliet, I. M., & van Veen, T. (2014). A comparative analysis of personality pathology profiles among patients with pure depressive-, pure anxiety-, and pure somatoform disorders. *Journal of affective disorders, 168,* 322-330.

Chambless, D. L., Renneberg, B., Fydrich, T., Goldstein, A. J., & Gracely, E. J. (1995). Axis I and axis II comorbidity in agoraphobic outpatients: prevalence and relationship to severity and treatment outcome. *Jounral of anxiety disorders, 6,* 193-211.

de Beurs, E., Barendregt, M., Rogmans, B., Robbers, S., van Geffen, M., van Aggelen-Gerrits, M., & Houben, H. (2015). Denoting treatment outcome in child and adolescent psychiatry: a comparison of continuous and categorical outcomes. *European child & adolescent psychiatry, 24,* 553-563.

de Beurs, E., den Hollander-Gijsman, M. E., van Rood, Y. R., Van der Wee, N. J., Giltay, E. J., van Noorden, M. S., ... & Zitman, F. G. (2011). Routine outcome monitoring in the

Netherlands: practical experiences with a web-based strategy for the assessment of treatment outcome in clinical practice. *Clinical psychology & psychotherapy, 18*, 1-12.

de Beurs, E., Rinne, T., van Kampen, D., Verheul, R., & Andrea, H. (2009). Reliability and validity of the Dutch Dimensional Assessment of Personality Pathology-Short Form (DAPP-SF), a shortened version of the DAPP-Basic Questionnaire. *Journal of Personality Disorders*, *23*, 308-326.

de Beurs, E., Rinne, T., van Kampen, D., Verheul, R., & Andrea, H. (2010). Criterion-related validity of the DAPP-SF and its utility as a screener for personality disorders in outpatient care. *Personality and Mental Health, 4*, 271-283.

de Beurs, E., & Zitman, F. (2005). De Brief Symptom Inventory (BSI). *De betrouwbaarheid en validiteit van een handzaam alternatief voor de SCL-90.* Leiden: Leids universitair medisch centrum.

de Vocht, A. (2013). *Basishandboek SPSS 21: IBM SPSS STATISTICS.* Utrecht: Bijleveld Press.

Derogatis, L. R., & Melisaratos, N. (1983). The brief symptom inventory: an introductory report. *Psychological medicine, 13*, 595-605.

Dreessen, L., & Arntz, A. (1998). The impact of personality disorders on treatment outcome of anxiety disorders: Best-evidence synthesis. *Behaviour Research and Therapy, 36*, 483-504.

Dreessen, L., Hoekstra, R., & Arntz, A. (1997). The influence of personality disorders on cognitive behavioural therapy for obsessive compulsive disorder. *Journal of Anxiety Disorders, 11*, 503-521.

Eurelings-Bontekoe, E. H. M., Verheul, R., & Snellen, W. M. (2009). *Handboek persoonlijkheidspathologie*. Houten: Bohn Stafleu van Loghum.

Faber, R. J., & O'Guinn, T. (1989). Classifying compulsive consumers: advances in the development of a diagnostic tool. *Advances in Consumer Research, 16*, 738-744.

Goddard, E., Wingrove, J., & Moran, P. (2015). The impact of comorbid personality difficulties on response to IAPT treatment for depression and anxiety. *Behaviour Research and Therapy, 73*, 1-7.

Harte, C. B., & Hawkins II, R. C. (2016). Impact of personality disorder comorbidity on cognitive-behavioral therapy outcome for mood and anxiety disorders: results from a university training clinic. *Research in Psychotherapy: Psychopathology, Process and Outcome*, *19*, 124-135.

Hoffart, A., & Martinsen, E. W. (1993). The effect of personality disorders and anxious-depressive comorbidity on outcome in patients with unipolar depression and with panic disorder and agoraphobia. *Journal of Personality Disorders, 7*, 304-311.

Illardi, S., Craighead, W., & Evans, D. D. (1997). Modelling relapse in unipolar depression: the effects of dysfunctional cognitions and personality disorders. *Journal of Consulting and Clinical Psychology*, *65*, 381-391.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of consulting and clinical psychology*, *59*, 12.

Joseph, B., Clayton, B., Jack, F. S., Kung-Yee, L., Paul, T. C., William, W. E., Nestadt, G. (2001). Normal personality traits and comorbidity among phobic, panic and major depressive disorders. *Psychiatry Research, 102*, 73–85.

Judd, L. L., Schettler, P. J., Coryell, W., Akiskal, H. S., & Fiedorowicz, J. G. (2013). Overt irritability/anger in unipolar major depressive episodes: past and current characteristics and implications for long-term course. *JAMA Psychiatry, 70*, 1171-1780.

Kampman, M., Keijsers, G. P. J., Hoogduin, C. A. L., & Hendriks, G. (2007). Outcome Prediction of Cognitive Behaviour Therapy for Panic Disorder: Initial Symptom Severity is Predictive for Treatment Outcome, Comorbid Anxiety or Depressive Disorder, Cluster C Personality Disorders and Initial Motivation Are Not. *Behavioural and Cognitive Psychotherapy, 36*, 99-112.

Lecrubier, Y., Sheehan, D. V., Weiller, E., Amorim, P., Bonora, I., Sheehan, K. H., ... & Dunbar, G. C. (1997). The Mini International Neuropsychiatric Interview (MINI). A short diagnostic structured interview: reliability and validity according to the CIDI. *European psychiatry*, *12*, 224-231.

Lewinsohn, P. M., Zinbarg, R., Seeley, J. R., Lewinsohn, M., & Sack, W. H. (1997). Lifetime comorbidity among anxiety disorders and between anxiety disorders and other mental disorders in adolescents. *Journal of Anxiety disorders*, *11*, 377-394.

Livesley, W. J., Jang, K. L., Jackson, D. N., & Vernon, P. A. (1993). Genetic and environmental contributions to dimensions of personality disorder. *The American Journal of Psychiatry*, *150*, 1826.

Mulder, R. T. (2002). Personality pathology and treatment outcome in major depression: a review. American Journal of Psychiatry, 159, 359-371.

Newton-Howes, G., Tyrer, P., Mulder, R., & Dekker, J. (2014). Influence of personality on the outcome of treatment in depression: systematic review and meta-analysis. *Journal of Personality Disorders*, *28*, 577-593.

Noyes, R., Reich, J., Christiansen, J., Suelzer, M., Pfohl, B., & Coryell, W. A. (1990). Outcome of panic disorder. Relationship to diagnostic subtypes and comorbidity. *Archives of General Psychiatry*, *47*, 809-818.

Pituch, K. A., & Stevens, J. P. (2016). *Applied Multivariate Statistics for the Social Sciences*: *analysis with SAS and IBM's SPSS (6th edition)*. New York: Routledge.

Reich, J. (2003). The effect of axis II disorders on the outcome of treatment of anxiety and unipolar depressive disorders: a review. *Journal of Personality Disorders, 17*, 387-405.

Reich, J. (2007). State and trait in personality disorders. *Annals of Clinical Psychiatry, 19*, 37-44.

Reich, J., & Vasile, R. (1993). Effect of personality disorders on the treatment outcome of axis I conditions: an update. *Journal of Nervous and Mental disease, 181*, 475-484.

Sanatinia, R., Wang, D., Tyrer, P., Tyrer, H., Crawford, M., Cooper, S. … Barrett, B. (2016). Impact of personality status on the outcomes and cost of cognitive-behavioural therapy for health anxiety. *The British Journal of Psychiatry*, *209*, 244-250.

Schat, A., van Noorden, M. S., Noom, M. J., Giltay, E. J., van Der Wee, N. J. A., Vermeiren, R. R. J. M., & Zitman, F. G. (2013). Predictors of outcome in outpatients with anxiety disorders: The Leiden routine outcome monitoring study. *Journal of Psychiatric Research, 47*, 1876-1885.

Seivewright, H., Tyrer, P., & Johnson, T. (1998). Prediction of outcome in neurotic disorder: A 5 year prospective study. *Psychological Medicine, 28*, 1149-1157.

Shea, M. T., Pilkonis, P. A., Beckham, E., Collins, J. F., Elkin, I., Sotsky, S. M., & Docherty, J. P. (1990). Personality Disorders and Treatment Outcome in the NIMH Treatment of Depression Collaborative Research Program. *The American Journal of Psychiatry, 147*, 711-718.

Siever, L. J., & Davis, K. L. (1991). A psychobiological perspective on the personality disorders. *American journal of Psychiatry, 148*, 1647-1658.

Skodol, A. E., & Bender, D. S. (2009). The future of personality disorders in DSM-V? *American Journal of Psychiatry, 166*, 388-391.

Steketee, G. (1990). Personality Traits and Disorders in Obsessive-Compulsives. *Journal of Anxiety Disorders, 4*, 351-364.

Telch, M. J., Kamphuis, J. H., & Schmidt, N. B. (2011). The effects of comorbid personality disorders on cognitive behavioral treatment for panic disorder. *Journal of Psychiatric Research*, *45*, 469-474.

Tyrer, P. Seivewright, N., Ferguson, B., Murphy, S., & Johnson, L. A. (1993). The Nottingham study of neurotic disorder: Effect of personality status on response to drug treatment, cognitive therapy and self-help over two years. *Journal of Psychiatry, 162*, 219-226.

van den Hout, M., Brouwers, C., & Oomen, J. (2006). Clinically diagnosed axis II co-morbidity and the short-term outcome of CBT for axis I disorders. *Clinical Psychology & Psychotherapy*, *13*, 56-63.

van Kampen, D., de Beurs, E., & Andrea, H. (2008). A short form of the Dimensional Assessment of Personality Pathology-Basic Questionnaire (DAPP-BQ): the DAPP-SF. *Psychiatry Research, 160,* 115-128.

van Vliet, I. M., & de Beurs, E. (2007). Het Mini Internationaal Neuropsychiatrisch Interview (MINI). Een kort gestructureerd diagnostisch psychiatrisch interview voor DSM-IV en ICD-10-stoornissen [The Mini International Neuropsychiatric Interview (MINI). A short structured diagnostic psychiatric interview for DSM-IV and ICD-10 disorders]. *Tijdschrift voor Psychiatrie*, *49*, 393-7.

Widiger, T. (1991). Definition, diagnosis, and differentiation. *Journal of Personality Disorders, 5,* 42-51.

Widiger, T. A., & Costa, P. T. (2012). Integrating normal and abnormal personality structure: the Five-Factor Model. *Journal of Personality Disorders, 80,* 1471-1506.

Appendix

Appendix 1

Frequencies of specific mood and anxiety disorders within the final data sample of N = 5755.

Clinical data	N	%
Major depressive disorder, recurrent, moderate	1232	21,4%
Posttraumatic stress disorder	898	15,6%
Generalized anxiety disorder	898	15,6%
Social anxiety disorder	675	11,7%
Panic disorder with agoraphobia	651	11,3%
Specific phobia	625	10,9%
Major depressive disorder, single episode, moderate	556	9,7%
Obsessive-compulsive disorder	470	8,2%
Major depressive disorder, recurrent, unspecified	438	7,6%
Dysthymic disorder	427	7,4%
Major depressive disorder, recurrent, severe without psychotic features	381	6,6%
Panic disorder without agoraphobia	301	5,2%
Major depressive disorder, recurrent, mild	272	4,7%
Major depressive disorder, single episode, unspecified	214	3,7%
Major depressive disorder, single episode, severe without psychotic features	203	3,5%
Agoraphobia without history of panic disorder	191	3,3%
Major depressive disorder, single episode, mild	179	3,1%
Major depressive disorder, recurrent, in partial remission	83	1,4%
Major depressive disorder, single episode, in	80	1,4%

partial remission		
Major depressive disorder, recurrent, severe with psychotic features	25	0,4%
Catatonic disorder due to general medical condition	17	0,3%
Major depressive disorder, single episode, severe with psychotic features	16	0,3%
Major depressive disorder, recurrent, unspecified	8	0,1%
Bipolar II disorder	6	0,1%
Bipolar I disorder, single manic episode, in partial remission	4	0,07%
Mood disorder due to general medical condition	4	0,07%
Anxiety disorder not otherwise specified	2	0,03%

Appendix 2

Multiple regression analysis examining the prognostic value of unique subscales of higher-order DAPP constructs for BSI after max. six months of treatment.

	B (SE)	Beta	*t*	*P*	95% CI
Selection variable: depression group					
Emotional dysregulation: $F(14, 1657) = 43.95$, $p <$ 0.05; Adj. $R^2 = 0.265$					
Submissiveness	-0.070(0.022)	-0.096	-3.183	0.001**	-0.114 to -0.027
Cognitive distortion	0.056(0.021)	0.078	2.645	0.008**	0.014-0.097
Identity problems	0.024(0.023)	0.034	1.034	0.298	-0.022 to 0.07
Affective lability	0.069(0.026)	0.088	2.678	0.007**	0.019-0.120
Oppositionality	0.065(0.021)	0.087	3.164	0.002**	0.025-0.106

Anxiousness	-0.036(0.024)	-0.050	-1.484	0.138	-0.083 to 0.012
Suspiciousness	0.050(0.020)	0.072	2.482	0.013**	0.01-0.089
Social avoidance	0,007(0.021)	0.011	0.333	0.739	-0.034 to 0.047
Narcissism	-0.032(0.020)	-0.040	-1.598	0.110	-0.071 to 0.007
Insecure attachment	-0.017(0.015)	-0.029	-1.131	0.258	-0.046 to 0.012
Self-harm	0.015(0.016)	0.022	0.916	0.360	-0.017 to 0.046
Dissocial behaviour: $F(14, 1664) = 79.31$, $p < 0.05$; Adj. $R^2 = 0.25$					
Stimulus seeking	0.001(0.022)	0.002	0.063	0.950	-0.041 to 0.044
Callousness	0.010(0.030)	0.009	0.335	0.738	-0.049 to 0.069
Rejection	-0.008(0.020)	-0.010	-0.385	0.700	-0.048 to 0.032
Conduct problems	0.089(0.031)	0.081	2.884	0.004**	0.028-0.149
Inhibitedness: $F(5, 1668) = 110.13$, $p < 0.05$; Adj. $R^2 = 0.25$					
Intimacy problems	0.067(0.017)	0.089	3.943	0.000**	0.034-0.100
Restricted expression	-0.004(0.025)	0.041	-0.214	0.830	-0.039 to 0.031
Compulsivity $F(4, 1674) = 136.25$, $p < 0.05$; Adj. $R^2 = 0.24$					
Compulsivity	-0.020(0.016)	-0.028	-1.298	0.195	-0.051 to 0.010
Selection variable: anxiety group					
Emotional dysregulation: $F(14, 1986) = 60.43$, $p < 0.05$; Adj. $R^2 = 0.29$					
Submissiveness	-0.050(0.017)	-0.082	-2.894	0.004**	-0.083 to -0.016
Cognitive distortion	0.050(0.017)	0.077	2.890	0.004**	0.016-0.084
Identity problems	0.015(0.019)	0.025	0.773	0.440	-0.023 to 0.052

Affective lability	0.024(0.020)	0.036	1.191	0.234	-0.015 to 0.063
Oppositionality	0.045(0.017)	0.067	2.589	0.010**	0.011-0.080
Anxiousness	0.024(0.019)	0.040	1.247	0.213	-0.014 to 0.063
Suspiciousness	0.052(0.017)	0.080	3.082	0.002**	0.019-0.084
Social avoidance	0.018(0.016)	0.034	1.147	0.251	-0.013 to 0.049
Narcissism	-0.009(0.016)	-0.013	-0.579	0.563	-0.041 to 0.022
Insecure attachment	0.004(0.012)	0.007	0.333	0.739	-0.019 to 0.027
Self-harm	0.064(0.017)	0.080	3.683	0.000**	0.030 to 0.098
Dissocial behaviour: $F(7, 1998) = 105.93$, $p < 0.05$; Adj. $R^2 = 0.27$					
Stimulus seeking	0.014(0.018)	0.019	0.768	0.442	-0.022 to 0.049
Callousness	0.070(0.024)	0.071	2.912	0.004**	0.023-0.118
Rejection	-0.027(0.016)	-0.038	-1.691	0.091	-0.057 to 0.004
Conduct problems	0.087(0.026)	0.080	3.302	0.001**	0.035-0.139
Inhibitedness: $F(5, 2002) = 144.77$, $p < 0.05$; Adj. $R^2 = 0.26$					
Intimacy problems	0.009(0.014)	0.012	0.596	0.551	-0.020 to 0.037
Restricted expression	0.059(0.014)	0.088	4.119	0.000**	0.031-0.087
Compulsivity $F(4, 2007) = 173.99$, $p < 0.05$; Adj. $R^2 = 0.26$					
Compulsivity	0.005(0.012)	0.009	0.441	0.659	-0.018 to 0.029
Selection variable: comorbid anxiety and depression group					
Emotional dysregulation: $F(14, 2027) = 45.22$, $p < 0.05$; Adj. $R^2 = 0.233$					
Submissiveness	-0.060(0.025)	-0.067	-2.359	0.018**	-0.110 to -0.010

Cognitive distortion	-0.007(0.023)	-0.008	-0.306	0.760	-0.052 to 0.038
Identity problems	0.022(0.028)	0.024	0.769	0.442	-0.034 to 0.077
Affective lability	0.068(0.030)	0.066	2.251	0.024**	0.009-0.128
Oppositionality	-.001(0.024)	-0.001	-0.026	0.979	-0.048 to 0.047
Anxiousness	-0.006(0.030)	-0.007	-0.205	0.837	-0.064 to 0.052
Suspiciousness	0.078(0.021)	0.099	3.671	0.000**	0.036-0.120
Social avoidance	0.008(0.023)	0.010	0.359	0.720	-0.037 to 0.054
Narcissism	-0.077(0.023)	-0.077	-3.385	0.001**	-0.121 to -0.032
Insecure attachment	-0.006(0.017)	-0.007	-0.317	0.751	-0.040 to 0.029
Self-harm	0.060(0.018)	0.076	3.423	0.001**	0.026 to 0.095
Dissocial behaviour: $F(7, 2041) = 84.81$, $p < 0.05$; Adj. $R^2 = 0.222$					
Stimulus seeking	0.027(0.025)	0.027	1.081	0.280	-0.022 to 0.075
Callousness	-0.024(0.033)	-0.0118	-0.720	0.472	-0.090 to 0.042
Rejection	-0.069(0.023)	-0.064	-2.731	0.006**	-0.109 to -0.018
Conduct problems	0.112(0.035)	0.081	3.221	0.001**	0.044 to 0.180
Inhibitedness: $F(5, 2046) = 122.65$, $p < 0.05$; Adj. $R^2 = 0.23$					
Intimacy problems	0.100(0.019)	0.107	5.334	0.000**	0.063-0.137
Restricted expression	-0.006(0.021)	-0.006	-0.287	0.774	-0.047 to 0.035
Compulsivity $F(4, 2056) = 145.70$, $p < 0.05$; Adj. $R^2 = 0.22$					
Compulsivity	-0.029(0.017)	-0.034	-1.669	0.095	-0.062 to 0.005

Note: DAPP-SF denotes Dimensional Assessment of Personality Psychopathology-Short Form; BSI denotes Brief Symptom Inventory; CI denotes confidence interval

* adjusted for gender and age

** $p < 0.05$